SUMMARY OF THE U.S. EPA COLLOQUIUM ON A FRAMEWORK FOR HUMAN HEALTH RISK ASSESSMENT

Colloquium #2

Prepared for:

U.S. Environmental Protection Agency
Risk Assessment Forum
401 M Street SW.
Washington, DC 20460

Contract No. 68-D5-0028
Work Assignment No. 3-98

Prepared by:

Eastern Research Group
110 Hartwell Avenue
Lexington, MA 02421-3136

September 1998

NOTICE

This report was prepared by Eastern Research Group, Inc. (ERG), an EPA contractor, as a general record of discussions during the U.S. EPA Colloquium on a Framework for Human Health Risk Assessment (Colloquium #2). As requested by EPA, this report captures the main points and highlights of discussions held during plenary sessions. The report is not a complete record of all details discussed nor does it embellish, interpret, or enlarge upon matters that were incomplete or unclear.

CONTENTS

Page

SECTION ONE BACKGROUND 1-1

Developing a Framework for Human Health Risk Assessment 1-1

The September 1997 Colloquium 1-2

The June 1998 Colloquium 1-2

SECTION TWO OPENING PLENARY SESSION 2-1

Welcoming Remarks 2-1

Goals of the Human Health Risk Assessment Framework 2-1

Introduction to Case Studies and Colloquium Issues and Charge to Breakout Groups 2-2

Questions/Comments 2-3

SECTION THREE BREAKOUT GROUP DISCUSSIONS ON CASE-SPECIFIC QUESTIONS 3-1

Ethylene Thiourea 3-2

Ethylene Oxide 3-5

Trichloroethylene 3-7

Vinyl Acetate 3-13

SECTION FOUR FINAL PLENARY SESSION 4-1

Lessons Learned and Their Applications to the Development of a Human Health Risk Assessment Framework 4-1

Overview/Next Steps 4-4

APPENDICES

APPENDIX A - White Paper

APPENDIX B - Participant and Observer Lists

APPENDIX C - Case Studies

APPENDIX D - Charge to the Participants

APPENDIX E - General Questions for Plenary Session

APPENDIX F - Breakout Group Assignments

APPENDIX G - Agenda

SECTION ONE
BACKGROUND

Developing a Framework for Human Health Risk Assessment

The U.S. Environmental Protection Agency (EPA) has recognized the need to develop a framework for human health risk assessment that puts a perspective on the approaches in practice throughout the Agency. Current human health risk assessment approaches are largely endpoint driven. In its 1994 report entitled *Science and Judgment in Risk Assessment,* the National Research Council (NRC) noted the importance of an approach that is less fragmented, more consistent in application of similar concepts, and more holistic than endpoint-specific guidelines. Both the NRC and EPA's Science Advisory Board have raised a number of issues for both cancer and noncancer risk assessments that should be reconsidered in light of recent scientific progress. EPA has recognized the need to develop a more integrated approach. In response, the Agency's Risk Assessment Forum (RAF) has begun the long-term process of developing a framework for human health risk assessment.

The framework will be a communication piece that will lay out the scientific basis, principles, and policy choices underlying past and current risk assessment approaches and will provide recommendations for integrating/harmonizing risk assessment methodologies for all human health endpoints.

As an initial step in this process, the RAF formed a technical panel in April 1996. An Issues Group (Gary Kimmel and Vanessa Vu, co-chairs; Jane Caldwell; Richard Hill; and Ed Ohanian) was formed, and this group developed a white paper, entitled *Human Health Risk Assessment: Current Approaches and Future Directions,* to provide an overall perspective on the issue (see Appendix A). The RAF peer-reviewed the white paper in February 1997. Its purpose is to serve as a basis for further discussion on current and potential future risk assessment approaches. The paper highlights a number of issues regarding the Agency's risk assessment approaches and their scientific basis, primarily with respect to dose-response and hazard assessment. The paper discusses the scientific basis for cancer and noncancer risk assessment, including differences and similarities. It also identifies knowledge/information gaps and areas where more work is needed.

As part of the continuing effort to develop a human health risk assessment framework, the RAF organized a colloquium series, consisting of two internal colloquia. The colloquia brought together EPA scientists for a dialogue on various scientific and policy issues pertaining to EPA's cancer and noncancer risk assessment approaches. The first colloquium, held on September 28 and 29, 1997, in Arlington, Virginia, focused on the role of mode of action information in re-examining and developing new risk assessment approaches. The second colloquium, held on June 2 and 3, 1998, in Bethesda, Maryland, explored the more quantitative aspects of mode of action, including dosimetry, dose-response relationships, and low-dose extrapolation methods.

The overall goal of the first two colloquia was to provide Agency scientists an opportunity to share perspectives on the role of mode of action in shaping future human health risk assessment approaches. The RAF invited a cross-section of senior Agency scientists (from headquarters, Research Triangle Park, Cincinnati, Las Vegas, and the regions) to participate in these discussions. As the Agency moves forward to develop this framework, additional colloquia are anticipated, as well as workshops to gather input and perspectives from scientists outside EPA.

The September 1997 Colloquium

During the first colloquium, Agency scientists discussed the current standard default approach for cancer and noncancer risk assessment, and the advantages and limitations of departing from this approach in light of new information pertaining to chemical mode of action. The primary topics deliberated by the group included defining mode of action, evaluating what events are critical, formulating dose metrics, determining when enough information exists to support new risk assessment approaches, and strategizing on how mode of action information can be effectively and systematically used in low-dose extrapolations. Group discussions addressed general risk assessment issues and the overall use of mode of action in risk assessment. Case study discussions followed. The colloquium's final session included discussions on "critical harmonization issues" and quantitative dose-response issues to be covered at the second colloquium.

The "Summary of the U.S. EPA Colloquium on a Framework for Human Health Risk Assessment: Colloquium #1," dated November 24, 1997, provides a detailed account of the outcome of the first colloquium. A brief overview of the key results of the September 1997 colloquium was provided at the opening session of the second colloquium (see Section Two).

The June 1998 Colloquium

Fifty EPA scientists and a small group of observers gathered for the second colloquium in June 1998 (see participant and observer lists in Appendix B). The 2-day colloquium focused on the role of mode of action information in developing descriptive quantitative models, applicable to a variety of needs for carrying out a risk assessment. Mode of action and harmonization issues were discussed in the context of four chemical-specific case studies: ethylene thiourea, ethylene oxide, trichloroethylene, and vinyl acetate.

Prior to the June colloquium each participant received one of the four case studies (Appendix C), including case-specific questions; a "charge" (Appendix D); and a list of general questions developed to guide colloquium discussions (Appendix E). During the colloquium, each participant was assigned to a breakout group to discuss assigned case studies. Appendix F includes a list of breakout group assignments, including the names of breakout group chairs and rapporteurs. As with the first colloquium, the RAF sought to ensure a mix of expertise and Agency representation in making group assignments.

After opening remarks were made, the first day of the colloquium was devoted to breakout group discussions on the case studies. During the second day, in plenary session, breakout group members presented their key findings. The closing plenary session involved an exchange of ideas on lessons learned from the colloquia series. Participants discussed next steps in developing a risk assessment framework in light of uncertainties and data gaps. The colloquium agenda is provided in Appendix G.

The following sections of this report highlight the outcome of the June 1998 colloquium. Section Two presents opening statements. Section Three captures the breakout group discussions on the case studies and Section Four presents highlights of the closing plenary session.

SECTION TWO
OPENING PLENARY SESSION

Welcoming Remarks
William Wood, Risk Assessment Forum, EPA

Dr. Wood welcomed all participants, many of whom were at the first colloquium. He explained that this RAF project was directed at developing a framework on integrating approaches for cancer and noncancer risk assessment. Toward that end, the RAF workgroup's goal is to couple the outcome of the health effects colloquia series with Agency work on the final cancer guidelines in setting the course for how EPA will conduct future risk assessments. The outcome of the colloquium will also provide guidance for future research. Dr. Wood encouraged the input and active participation of Agency scientists throughout the second colloquium.

Dr. Wood acknowledged the hard work of the organizing committee whose members include Gary Kimmel (co-chair), Vanessa Vu (co-chair), Kim Hoang, Annie Jarabek, Jennifer Seed, Gina Pastino, and Wendy Yap. The colloquium participants then introduced themselves and their affiliations.

Goals of the Human Health Risk Assessment Framework
Vanessa Vu, National Center for Environmental Assessment, EPA

Dr. Vu reviewed the overall goals of the framework project, accomplishments to date, additional short- and long-term plans, and the structure and charge of the second colloquium. She explained that the Agency intends to develop a framework to accomplish the following:

- Develop a conceptual piece to communicate a risk assessment approach (for the Agency and public at large).
- Layout past and current approaches.
- Recommend approaches in integrating/harmonizing risk assessment approaches for all endpoints.

The major elements of the anticipated framework, she explained, include using mechanistic information to enable integrating risk approaches for different endpoints, considering a range of default approaches, and applying appropriate uncertainty factors.

Obtaining buy-in and input from Agency scientists, Dr. Vu emphasized, is very important, especially in the development stage of the framework. The RAF, therefore, has or plans to take the following steps:

1. *Development of a white paper*. The white paper, a "perspective" piece, was developed to identify key issues related to current risk assessment approaches and harmonization. The papers focuses on issues related to hazard and dose-response assessment and presents the scientific basis for assessing cancer and noncancer risks. It identifies uncertainties in the existing risk assessment process and areas requiring further guidance and research.

2. *Organization of the colloquia series*. The RAF organized the colloquia series to enable Agency scientists to discuss white paper issues and to provide recommendations on the approach of the framework. Agency scientists participating in the colloquia series were charged with discussing scientific and policy issues associated with developing a more consistent/holistic approach to risk assessment. During the first colloquium, discussions centered on the significance of qualitative implications of mode of action for various risk assessment endpoints. The second colloquium was designed to foster further qualitative discussions and initiate discussions on quantitative issues associated with the application of mode of action information (e.g., low dose extrapolation models).

3. *Draft the framework*. Based on the outcome of the colloquia series, the Agency anticipates preparing a draft framework document. It is anticipated that the framework document will undergo expert review, leading to future workshops and review by the Science Advisory Board.

Dr. Vu briefly summarized the outcome of the first colloquium. During Colloquium #1, participants developed a common appreciation for terminology and the role of mode of action in risk assessment. While Colloquium #1 participants recognized that strictly defining mode of action was difficult, mode of action was broadly defined as "knowledge of the series or sequence of biological events that influence the final toxic outcome." The group agreed that the traditional use of threshold/nonthreshold approaches may no longer be applicable in light of new scientific knowledge on mode of action. The group recommended greater use of mode of action information when extrapolating from high to low dose, across species, and across routes of exposure, as well as studying aggregate risk from chemicals that may have common mode of action. Colloquium #1 case studies enabled participants to begin to explore new approaches to low-dose extrapolation and evaluate commonalities across endpoints by reviewing toxicologic and mechanistic information for five chemicals. Participants agreed that issues related to commonalities across toxicities needed more emphasis. Continued development of the framework and future colloquia/workshops were encouraged to pursue the complex issues associated with harmonization of risk assessment approaches.

Introduction to Case Studies and Colloquium #2 Issues and Charge to Breakout Groups

Dr. Vu explained that the purpose of the case study exercise at the second colloquium was to foster more in depth discussions on critical issues related to mode of action and its role in harmonizing

cancer/noncancer risk assessment. Dr. Vu emphasized that the intent of the case studies was not to perform chemical-specific risk assessments. Dose-response and mechanistic data were provided to help participants explore important factors related to developing descriptive quantitative models. Case-specific questions were provided to guide discussions and to promote deliberations on harmonization issues.

Lastly, Dr. Vu acknowledged the efforts of the issues group, organizing committee, RAF (Bill Wood, Jeanette Wilsey, and Carole Kimmel), and Eastern Research Group in helping to organize and coordinate the activities of the workshop. Dr. Vu also thanked participants and observers for taking part in the colloquia series.

Questions/Comments

The group briefly discussed possible limitations of the case studies. Points raised by participants include the following:

- Chemical-specific information presented in the case studies may not be 100 percent complete or correct. One participant questioned whether discussions should be limited to information provided in the case studies or if new information could be introduced.

  The group recognized that it would be impossible to present a complete data set for one- or two-day discussions on a particular chemical. It was re-emphasized that participants were not performing full-blown risk assessments on case-study chemicals, but rather raising and evaluating case-specific issues related to more scientifically sound approaches to evaluating human health risks. While it was agreed that scientists should introduce pertinent data during the breakout sessions, it was also recognized that because of time constraints it is not possible, nor necessary, to consider every chemical-specific detail. The ultimate purpose of the case study exercise, the group was reminded, was to determine the best use of mode of action information and how to generate the most credible risk assessment.

- One participant questioned how the group should approach the issue of multiple modes of action during case study deliberations, expressing concern that the group may try to "force fit" a single mode of action for multiple endpoints.

  Multiple modes of action should be considered in terms of their relative contribution to pathogenesis. The intent of the case study exercise was to evaluate whether different endpoints should be treated differently when a common mode of action has been identified, not necessarily to identify a single mode of action.

SECTION THREE
BREAKOUT GROUP DISCUSSIONS ON CASE-SPECIFIC QUESTIONS

The first day of the colloquium was dedicated to breakout group discussions on the following four case studies (See Appendix C).

- # Ethylene Thiourea (ETU)
- # Ethylene Oxide (EtO)
- # Trichoroethylene (TCE)
- # Vinyl Acetate (VA)

The case studies include a summary of key human and animal studies and describe primary acute and chronic effects. Depending on the chemical, the case study describes portal-of-entry effects; systemic toxicity; reproductive and developmental toxicity; neurotoxicity; mutagenicity; and carcinogenicity. The case studies also present pertinent dose-response, pharmacokinetic, and mode of action (MOA) information.

Each breakout group deliberated case-specific questions (included within each case study), but, in general, the following questions capture the key issues discussed by each group.

1. Given what is known about MOA, are there commonalities among endpoints that would be useful for quantitative analyses? For which endpoints should a common quantitative analysis be conducted? For which endpoints should a separate analysis be conducted?

2. What additional information would be useful for quantitative analysis?

3. In the absence of this information, are any of the available data sets useful for quantitative analysis?

4. Are dose and duration of exposure important considerations? If so, for which endpoints and how should they be handled?

5. In the absence of case-specific physiologically-based pharmacokinetic (PBPK) models, how should dose be adjusted for extrapolation to humans? Does choice of a specific endpoint influence this decision? Why or why not?

   If a PBPK model is available, which dose metrics should be considered for the dose-response analysis?

6. What response/endpoint(s) would be useful for dose-response modeling in the observable range? Does MOA information influence this choice?

7. What quantitative method is recommended for low level exposures? Does this vary for different toxicities? Does MOA information influence the choice of models?

8. If a reference dose (RfD), reference concentration (RfC), or margin of exposure (MOE) were to be calculated, does MOA information influence the choice of uncertainty factors or influence uncertainties about data gaps?

The sections below summarize the main points discussed during breakout sessions, as captured by the group rapporteurs and presented in plenary session. Vicki Dellarco, Kerry Deerfield, Vanessa Vu, and Arnold Kuzmack presented the breakout group reports for ETU, EtO, TCE, and VA, respectively.

Ethylene Thiourea

In reviewing the ETU case study, the group considered the adverse health effects associated with target organs/responses, common modes of actions across different responses, dose-related increases, exposure duration issues, critical windows of exposure, and the relevancy of animal data to humans. The group's responses to case-specific questions are provided below.

Given what is known about MOA, are there commonalities among endpoints?

The group identified the following ETU "targets:" thyroid, pituitary, liver, embryo/postnatal, and central nervous system (CNS). The group described the following three potential modes of action likely to be responsible for the effects in these target systems:

1. Thyroid/pituitary: In the rat, high concentrations of ETU result in decreased T3 and T4 and increased TSH levels. The severity of hyperplasia increases with dose and possibly with duration. These changes in T3/T4 and TSH levels are associated with thyroid hyperplasia and tumor development in the thyroid (adenomas and carcinoma). These events can eventually lead to pituitary tumors if substantial. Based on case study information, mutagenicity or a direct DNA reactive mechanism does not seem to be a major influence on tumor development. Perturbances of the pituitary-thyroid homeostasis is the essential event leading to tumor development (i.e., an anti-thyroid MOA).

   Some developmental effects (related to brain development in late gestational/postnatal periods) are presumed to by thyroid-mediated.

2. Liver: A separate MOA appears responsible for liver effects. Effects appear to be metabolite-dependent (FMO) and species-specific.

3. Non-thyroid developmental effects: Seen primarily in the rat, CNS malformations result from necrosis of neuroblasts driven by ETU (parent compound). These effects are not considered to be thyroid-mediated. Effects are species-specific.

The group concluded that common modes of action for cancer and certain noncancer (e.g., CNS) endpoints are associated with the disruption of the thryoid/pituitary homeostasis. This knowledge enables one to use precursor events (e.g., changes in T3, T4, and TSH; increases in thyroid hyperplasia) instead of frank toxicologic effects in protecting for different outcomes. No conclusions could be reached on the reversibility of responses, however, because of the lack of data.

A question was raised following this discussion as to whether or not ETU exposures led to total endocrine disruption and whether the pituitary should be considered separately. Another participant questioned whether the group considered the relation of liver effects to thyroid/pituitary effects. It was noted that data were not available to suggest any such relation.

What approaches should be considered for quantitative analysis?

Upon consideration of available dose-response data, the group suggested different approaches for the quantitative analyses of the three identified MOAs. For thyroid/pituitary events and hyperplasia events, effects on thyroid hormones should be used as indicators of both cancer and noncancer endpoints. Given the understanding of MOA in the thyroid, the group suggested using a nonlinear approach for low-dose extrapolation. For liver effects, the group noted that, in the absence of quantitative information and a full understanding of MOA, the default linear approach should be used. The group commented, however, that this approach might be overly conservative—the group emphasized the need to point out data set uncertainties and the possibility that effects may be species-specific and not relevant to humans. For developmental effects, the group suggested using the default nonlinear approach, but data were available to also enable some benchmark modeling.

What additional information would be useful for quantitative analysis? What are the research needs?

In general, the breakout group agreed that more comparative metabolism information (within and across species) would be especially helpful in further evaluating MOA questions and the relevance of existing data to humans. Response-specific information needs to include the following:

Thyroid: Because thyroid hormones are a good biomarker and evidence exists that there is age-dependent susceptibility, it would be helpful to examine prenatal/early postnatal hormone levels. In addition, obtaining more dose-duration information would be helpful in studying the issue of reversibility. Comparative metabolism data (tissue distribution) between humans and rodents would be helpful to better understand species differences.

Liver: More information is needed specific to mouse metabolism. Comparative metabolism studies on FMO are needed.

Nonthyroid Malformation: More comparative metabolism data are needed to study differences in responses between humans and rats.

Are dose and duration exposure important considerations?

The breakout group considered patterns of exposure and critical windows of susceptibility. Responses in the thyroid/pituitary (severity of hyperplasia) appear to be dose limited and may be dependent on duration. Not enough information is available to assess dose/duration considerations for liver and developmental effects. Thyroid/pituitary and developmental effects were observed at similar ETU doses. Dose was species-dependent for liver effects, which is an example of why more species-specific metabolism data are needed.

In the absence of a PBPK-model, how should dose be adjusted for extrapolation to humans? Does choice of a specific endpoint influence this decision? What quantitative method is recommended for low level exposures? Does this vary for different toxicities? Does MOA information influence the choice of models?

Although no single extrapolation method was recommended (e.g., lack of an interspecies adjustment versus using a scaling factor of body weight to the 3/4 power), the group strongly agreed that the approach should be the same for cancer and noncancer endpoints in the thyroid/pituitary.

What endpoint(s) would be useful for dose-response modeling in the observable range? Does MOA information influence this choice?

The group agreed that MOA is relevant to thyroid/pituitary responses. It plays less of a role in developing models for liver and developmental effects.

If an RfD were to be calculated, does MOA information influence choice of uncertainty factors or influence uncertainties about data gaps?

Yes. The group reiterated, however, that more comparative data between rats and humans are needed before fully answering this question. Qualitatively, the group agreed that uncertainty factors should be applied in the same way for cancer and noncancer endpoints. In comparing RfD and margin of exposure (MOE) approaches, the group agreed that, conceptually, the uncertainty factors applied are similar. In practice, however, they could be applied differently because the RfD approach is more compartmentalized and the MOE approach involves more scientific judgment/interpretation. This issue, therefore, warrants further study and careful consideration.

Ethylene Oxide

The breakout group initiated their evaluation of EtO by preparing a matrix of observed effects. EtO induces a variety of effects including irritation, hematoxicity, neurotoxicity, reproductive and developmental toxicity, and cancer. Group discussions focused primarily on the latter three. The group provided the following responses to case-specific questions:

Given what is known about MOA, are there commonalities among toxicities that would be useful for quantitative analyses? Is there any reason to propose different mechanisms for the various endpoints?

Based on available data, two plausible MOAs exist for EtO: the formation of protein adducts and the formation of DNA adducts. EtO distributes readily and is direct acting (no metabolite formation). Distribution is even throughout the body. Although it is highly reactive (e.g., hemoglobin binding, glutathione binding), free EtO distributes to target tissues. EtO binds to macromolecules (specific amino acids in protein) and forms specific DNA adducts (e.g, 7-hydroxyethylquanine). These two mechanisms are probably not mutually exclusive. The mechanisms related to neurotoxic outcomes are not completely understood; these effects are not fully explained by DNA adduct formation, and may relate primarily to the binding of EtO to protein.

The group categorized the endpoints and asked whether common MOAs exist.

Cancer: Tumors have been observed in multiple sites in animals (hematopoietic, brain, forestomach, lung, ovary, lymph). In humans, epidemiologic studies suggest a link between EtO and hemopoietic cancers. Because tumors appear in multiple locations, there is likely a common MOA for most of these cancers and that is related to DNA binding mechanisms. Forestomach cancers, however, appear to result from a local irritant effect, although this effect may be enhanced by the genotoxic action of EtO.

Reproductive/Developmental Effects: Observed effects include spontaneous abortion, zygotic death, lethality/viability, litter size, implant loss, and malformations. Dominant lethality appears to result from the formation of DNA adducts. While insufficient data exist for all of these endpoints, the group agreed that a common MOA probably exists for most reproductive/developmental endpoints.

Data suggest that MOA is similar in animals and humans for tumors, but unknown for developmental effects.

What additional information would be useful for quantitative analysis of the various toxicities? (For example, is consideration of the entire spectrum of mutational changes, such as the induction of gene mutations, structural chromosome mutations, and numerical chromosome alterations important?)

Several data needs were identified.

- For mutagenic effects, existing information on point mutations needs to be considered. The case study concentrated on chromosome breaks (translocation) data.

- Information on the shape of the curve at low doses. For example, is it linear or nonlinear? Are DNA adducts formed at low levels? This is a research need.

- Additional information on the causality of different endpoints.

- Cell proliferation information at all dose levels.

- Information on background rates. What is the background load (endogenous EtO)?

- Information on exposures to other agents that may have the same MOA or make one more susceptible to a MOA.

What quantitative method is recommended for low level exposures? Does this vary for different responses? Does MOA information influence choice of models?

The group proposed the same approach for both cancer and developmental/reproductive effects because the MOA suggests that both effects are related to the formation of DNA adducts. If one assumes linear behavior, then a linear quantitative method is appropriate for low dose extrapolation because of the mutagenic properties of EtO. The group, however, did discuss MOE and possible nonlinear approaches because the data suggest that protein binding and DNA adduct formation may not be linear. One participant noted that data on heritable effects versus dominant lethal effects suggest that a two-hit model and nonlinear dose response may exist. The overall impression of the group was that MOE eliminates the theoretical argument over linear versus nonlinear dose-response relationships and focuses on MOA. MOE would therefore be a viable approach to bring to the risk manager. In general, the MOA for all effects is probably related specifically to the electrophilic nature of EtO, and the ultimate action would be dependent on timing and duration of exposure, where and to what it binds, etc.

The question on linear versus nonlinear dose response triggered a fairly lengthy discussion among the plenary group. General and EtO-specific issues raised are highlighted below:

- Because of the limited dose numbers in the NTP study, it is difficult to study linearity.

- Adduct-formation is not the only factor to influence the shape of the dose-response curve. Although adduct formation may be considered a linear response, a certain level may need to be reached before a toxic outcome is observed. If adducts are easily repaired, a nonlinear response may in fact be observed. What is happening beyond adduct formation needs to be considered and is an argument for using the MOE approach.

- "Toxicity" needs to be defined. Traditionally, toxicity was defined as an observable effect (e.g., a tumor or malformation). Now with activities at the cellular level being considered (e.g., biochemical changes or adduct formation), toxicologists need to agree on what the "toxic endpoint" is.

  One participant noted a definition of toxicity by Doull (of Cassarett and Doull): toxicity is

not achieved until the first "irreversible step" is observed. Several others disagreed citing ethanol exposure as an example where reversible effects still result in "toxicity." Furthermore, RfDs have been developed based on nontoxic reversible effects. Doull's definition, therefore, may not be relevant to these discussions.

- It is important to study the nature of the lesions before deciding on a linear versus nonlinear approach.

Are dose and duration of exposure important considerations? If so, which responses and how should they be handled?

Very little dose rate information is available for most endpoints, but the group agreed that it is an important consideration. For example, in a study of dominant lethality, dose and duration were found to be extremely important when considering the effects of EtO.

In summary, it was agreed that EtO presents a good case for quantitatively treating different endpoints similarly based on MOA. Although no specific approach was recommended, many felt that an integrated MOE approach for each of the effects would provide risk managers with useful information.

Trichloroethylene

TCE, the group agreed, was one of the more complex case studies because of the variety of systems affected and effects produced. It is further complicated because of the involvement of and uncertainties associated with the metabolites. The group reviewed TCE effects and its MOA in several target systems, but focused on effects in the liver, lung, and kidney.

Both the "minor" and "major" metabolic pathways for TCE were described (see case study figure in Appendix C). The group identified the role of metabolites in mediated TCE-induced toxicities and highlighted the relative species reactivity of the metabolites, as follows:

Effects	**Metabolites**	**Species reactivity**
liver	TCA, DCA	mouse>rat>humans
lung	Chloral	mouse>rat
kidney	DCVC	rat>mouse>human

TCA = trichloroacetic acid
DCA = dichloroacetic acid
DCVC = s-1,2-dichlorovinyl cysteine

The breakout group summarized the effects of TCE in the liver, lung, and kidney, highlighting cross-species and general dose duration differences. These discussions are summarized in Table 1.

Table 1. Breakout Group Summary of TCE Effects

SPECIES	EXPOSURE	EFFECTS
Liver		
human	acute/high	liver failure/necrosis
	occupational	impaired liver function
		some evidence of risk of cancer of the liver and the biliary duct
rat	acute/subchronic to high level	enlarged liver, hypertrophy, necrosis
	chronic/lower level	enlarged liver
mouse	acute/subchronic to high level	enlarged liver, hypertrophy, necrosis
	chronic/relatively lower level	hepatomegaly, hypertrophy, tumors
Kidney		
human	occupational	mild renal function changes suggestive evidence of kidney cancer
rat	acute exposure to high level	nephropathy
	chronic to lower level	increased kidney weight mild karyomegaly tumors
mouse	acute/chronic to high level	nephrotoxicity no tumors
Lung		
human		no reported effects
rat	acute/chronic	no effects
mouse	acute	cytotoxicity to Clara cells
	chronic	lung tumors

The group briefly discussed lympho/hematopoietic, reproductive/developmental, and CNS effects. TCE-related effects on the lympho/hematopoietic system include excess non-Hodgkin lymphoma in humans, lymphoma in exposed mice (via inhalation), and effects on the spleen in rats and mice. The group noted consistency across species. Inconclusive/conflicting evidence exists related to TCE-induced reproductive/developmental effects in humans. Eye and cardiac malformations have been observed in rats exposed *in utero*. Effects on sperm, implantations, and litter size have been observed in mouse reproductive studies. CNS effects are reported in humans exposed to high levels of TCE (acutely) and in occupational settings as well as in rats and mice exposed acutely, subchronically, and chronically.

Having highlighted key effects, the group then answered case-specific questions.

What seems to be the series of events leading to each observed toxic response? Are there any reversible steps in the process? Can an irreversible step be identified in each process? Given that TCE-induced toxicities are mediated through metabolites, are there common biological responses across toxicities that would be useful for quantitative analyses?

The group developed schematics depicting key events in the liver, kidney, and lung (see Figures 1, 2, and 3). Discussions centered around whether common modes of action are present for different toxic responses.

Figure 1. TCE MOA in the Liver

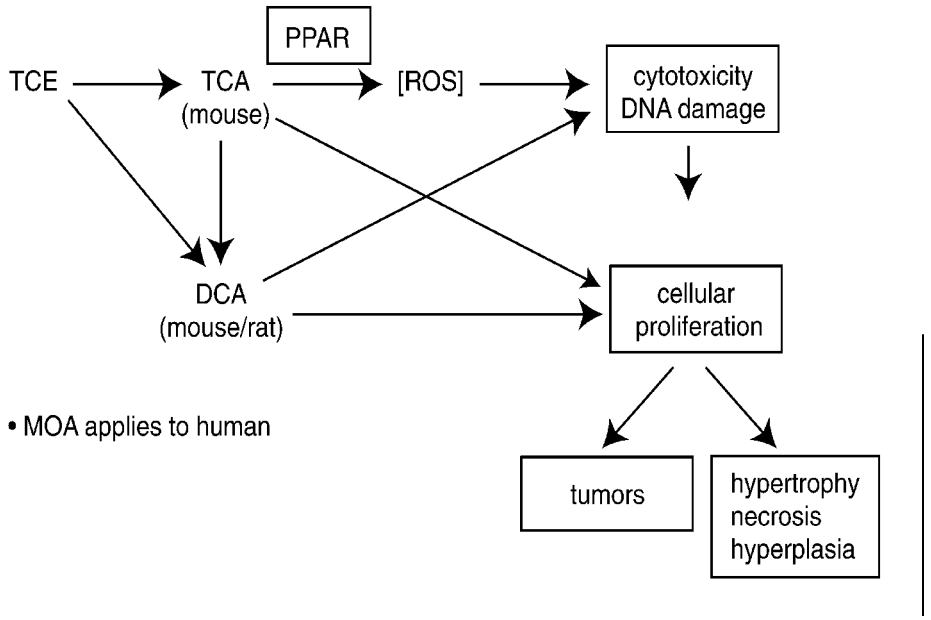


Figure 2. TCE MOA in the Kidney (Rat)

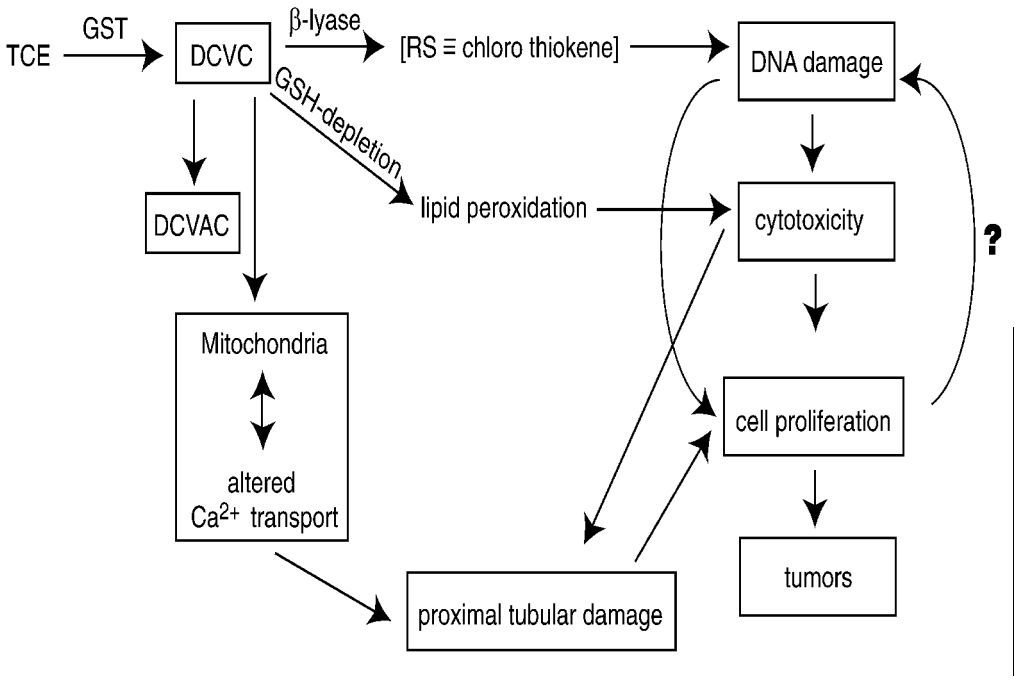
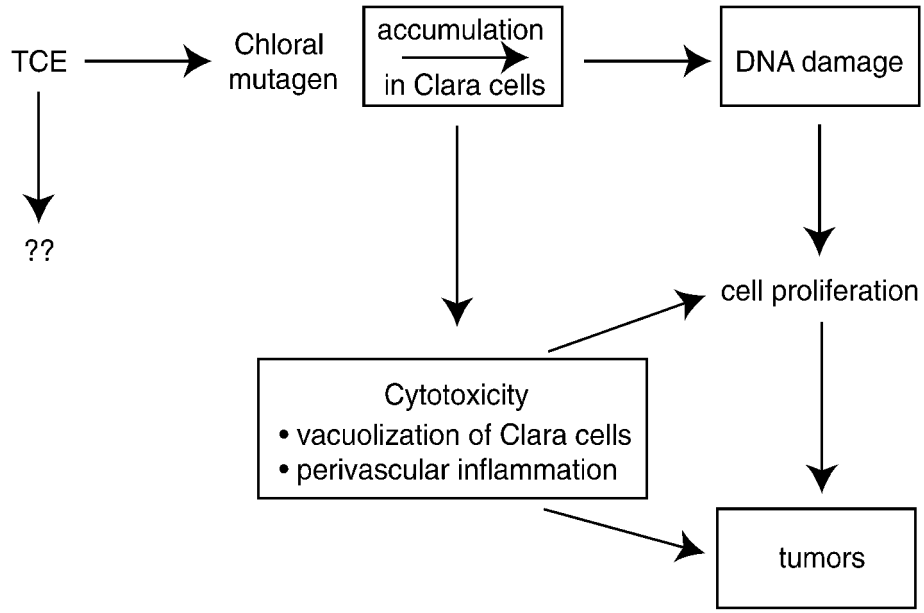


Figure 3. TCE MOA in the Lung (mouse)



The group also identified data gaps.

> *Liver:* For the liver, the group emphasized that cellular proliferation appears to be the common event leading to both tumors and liver toxicity. The MOA is relevant to humans based on available data. Because quantitative information on cellular proliferation is lacking, it is not known whether reversible steps exist. The specific steps leading to tumors and liver toxicity are not clear. One group member noted that speculation exists as to whether DCA is a promoter or an initiator.
>
> *Kidney*: While the metabolite DCVC is common to the two endpoints (i.e., tumors and proximal tubular damage), a common MOA is not observed for these endpoints.
>
> *Lung:* TCE action in the lung of mice was described. Both cytotoxicity and DNA damage appear to be the result of the accumulation of chloral in the Clara cells. Because of many unknowns, no specific common biological events could be identified to account for either TCE-induced tumors or toxicity in the lung.

Which of the above-selected responses is most relevant to humans regarding specificity (response concordance) and sensitivity (dose range of response)?

Liver and kidney MOA and responses in test animals are relevant to humans. Lung responses, however, are not. Data are not sufficient to judge sensitivity of response. Epidemiologic data provide good qualitative information but do not enable quantification. Animal studies show more tumors in the liver versus the kidney following TCE exposure.

What additional information would be useful for quantitative analysis?

The group stressed that obtaining more dose-response information on cell proliferation was critical. No dose-response curve is available. Cell proliferation data are needed for initiated versus noninitiated cells. A labeling index study for age range is also needed.

Are dose and duration of exposure important considerations? If so, for which toxicity and how should they be handled?

Dose and duration appear to be important in the liver and the kidney. In animal studies, liver tumor response depends on dose, but not enough is known to specifically answer the dose/duration question. Not enough data are available to answer this question for the kidney. In addition, more information is needed on dose/duration issues in humans.

What response(s) would be useful for dose-response modeling in the observable range for each toxicity? How does MOA information influence this choice? Given the availability of the PBPK models, what would be the appropriate dosimeters for the toxicity observed in the liver, lung, and kidney? Which quantitative models should be used for the observed data?

Dose-response modeling could be considered for liver and kidney responses. Cell proliferation in the liver is the preferable response choice, but because of high background and species variability, coupled with the lack of quantitative data, it may be problematic. PBPK models could be used to estimate internal TCE dose. More information is needed, however, relating TCE to its metabolites so that an internal dose of metabolites can be obtained.

Given what is known about the MOA for each toxicity, what quantitative approach would be recommended for characterizing risk associated with low level exposures (i.e., beyond the observable range) for each toxicity?

The group focused on the liver response for this question. Opinions varied regarding the best quantitative approach to take in light of available data. Although no one approach was recommended, it was agreed that applying a biologically-based dose response (BBDR) model would be the ideal choice. The group considered two scenarios: (1) assume quantitative cell proliferation data are available, and (2) assume quantitative cell proliferation data are not available.

Assuming quantitative cell proliferation data were available, the group considered linear and MOE approaches. Half of the breakout group felt an MOE approach was preferable because it gives more consideration to science and nearly an equal number felt it is really a policy choice. One individual preferred a linear approach because it is more conservative and because the threshold for lifetime exposure is not known.

In the absence of cell proliferation data or a BBDR model (where tumor and liver toxicity would be considered as the responses), the group was again divided as to what approach is most appropriate. The following quantitative approaches were proposed, with the group divided equally on each of the three options.

1. *Status quo*. Several individuals supported using default approaches (i.e., linear for tumor and an RfD/RfC for noncancer effects). These individuals felt resorting to the existing models was more conservative in light of data gaps.

2. *Same approach for both responses*. Because of common MOA, others felt it was more appropriate to use the same approach for both cancer and noncancer outcomes. Both linear and MOE approaches were considered. The overall preference of the group was an MOE approach because of observed receptor-threshold effects. One member noted that, in the absence of data, no compelling reason exists to assume a linear curve at low doses; he emphasized, however, that all endpoints should be considered and the most sensitive should be used to select the RfD/benchmark dose.

3. *Policy choice*.

Presentation of these choices resulted in lively discussions both in breakout and plenary sessions. The group conveyed the following general points about choosing an appropriate quantitative approach.

- **#** How much information is enough to support a decision to choose a nondefault approach? Because of the uncertainties in most data sets, opinion will vary widely.

- **#** In the case of TCE, one participant questioned how one could conclude simply from the evidence of cell proliferation whether a threshold or nonthreshold response existed. He provided the dioxin example where several factors led to identifying a threshold. He could not accept the threshold concept for the complicated TCE story.

- **#** The group did not discuss other sensitive noncancer effects of TCE (e.g., neurotoxic effects). In focusing on the noncancer effects in the liver (cell proliferation), a potentially more sensitive outcome in another system (neurotoxic) may be overlooked.

If an RfD or MOE were to be developed, which factors should be considered to account for uncertainties in risk assessment?

The group agreed that the following uncertainty factors should be considered as common to both RfD and MOE approaches:

- — intraspecies differences
- — interspecies differences
- — nature of response
- — steepness of the dose-response curve at point of departure region
- — lack of understanding

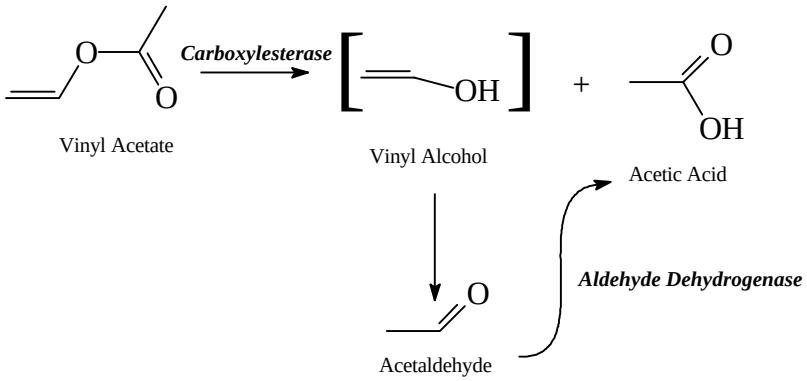
Further discussion on uncertainty factors was held in the final plenary session and is summarized in Section Four of this report.

Vinyl Acetate

It was noted that the action of VA is unique from the chemicals evaluated in the other case studies in that it exhibits effects at the portal-of-entry (upper respiratory tract). There is a spatial specificity of lesion location, with most effects concentrated in the olfactory region of the rat. In mice, the location of the lesions is consistent with air-flow patterns and tissue-specific enzymes. Case-specific questions varied slightly, therefore, to foster discussions on this unique aspect of VA.

The group reviewed the established metabolic pathway for VA. Carboxylesterase catalyzes the initial hydrolysis of VA to vinyl alcohol and acetic acid (AA). Vinyl alcohol rearranges to acetaldehyde (AAld) which aldehyde dehydrogenase subsequently metabolizes to additional AA. These enzymes have been localized histochemically and are found in discrete cell types in the respiratory and olfactory mucosae. The metabolism scheme (as presented in the case study) is depicted in Figure 4:

Figure 4. Metabolic Pathways for VA



Two mechanisms of action were identified: (1) AA causes cytotoxicity which may progress to cell proliferation, (2) AAld, which is a known clastogen and sister chromatid exchange initiator, leads to multi-hit genetic damage. Tumors are seen only in male rats at the highest concentration tested, 600 parts per million (ppm), and only at the terminal sacrifice of a 2-year bioassay; no effects are observed at concentrations below 50 ppm. It was hypothesized that because mice can restrict respiration (reflex apnea), less of an effect is observed. This species difference was shown to be the case with formaldehyde, another upper respiratory tract (URT) irritant.

Does the existing database support the URT lesions as the sentinel toxicity for inhalation exposures to VA?

The group agreed the database clearly supports URT lesions as the sentinel toxicity. The proximal to distal pattern and the concentration response are both important to the argument.

Can the cytotoxic changes caused by VA exposure be considered as sequentially linked to the observed tumor outcome? What are the key considerations to characterize the conditions of hazard (e.g., high dose versus low dose)? How do the genotoxic data factor in this characterization?

- **#** Cytotoxic changes caused by VA are linked to tumors.
- **#** AAld are linked with different tumor types. Responses in both pathways appear to be at high doses only. The group noted that the spatial distribution of tumors was consistent.
- **#** A "good" PBPK model exists that relates metabolism, physical layout, and fluid mechanics in human and rodents. The PBPK model accounts for the observed species and gender differences.
- **#** Knowledge of cytotoxicity, cell proliferation and temporal aspects, and localization of enzymes is helpful.

- Cytotoxicity may cause death but some cells will survive and those will have an increased probability for genotoxic effects, especially at high concentrations.

What mechanistic data are most relevant to characterizing tumor outcome? Which would be useful for dose-response modeling in the observable range? What are the implications of the MOA information for extrapolation of risk to low dose?

The data most relevant to tumor outcome include: cytotoxicity, cell proliferation, genotoxicity, site specificity (localization of effect), and metabolism. Dose-response modeling based on tumor outcome is not possible, however, because only two non-zero points (the second lowest with a response of 1) exist in the observable range. Because effects are seen only at the highest exposure concentration and only at the last sacrifice, the group overall felt this suggests that a nonlinear approach is appropriate for low-dose extrapolation. This was supported by clear relationships of genotoxicity, cytotoxicity, and cell proliferation only with high concentrations.

One breakout group member, however, disagreed that all effects are only at high concentrations. He noted that AA leads to cytotoxicity as a result of changes of pH, which may ultimately lead to cellular changes in the URT and to cancer. He agreed that the effect of AAld is significant only at high doses. Evidence includes the fact that cross links are only significant at high doses and that there are no long-lived DNA adducts. He noted, however, that large-scale changes in DNA have been observed that may have required multiple events. He noted that these large-scale changes are important to humans and should be examined closely. Dose-response data are lacking for observed DNA damage. In addition, there is a lack of mechanistic understanding of the process. A low dose linear situation may, therefore, exist.

Given the availability of the PBPK model, which dose metrics should be considered for the dose-response analysis? Does this choice of dose metric address consideration of the role of exposure duration?

Limited time was spent discussing the PBPK model although its usefulness in addressing the toxicokinetic issue of species to species extrapolation was recognized. The dose metrics (about seven tabulated) need to be further explored for implications to quantitative dose-response assessment. At 50 ppm VA, the model predicts the same decrement in pH projected in animals and humans. The group concluded that, at lower doses, animal and human responses would be quantitatively the same, but that the case study did not present the model in sufficient detail to quantitatively explore the interspecies differences in dosimetry (e.g., airflow).

What are the uncertainties in using these data to characterize human risk?

The group identified several uncertainties and data gaps that, if filled, would enable further consideration of the mechanistic actions and commonalities across endpoints.

- Reflex apnea in mice.

- Description of lesions (coverage in case study was brief).

- Effects of lowered pH in the respiratory tract on cancer.
- Effects of acetic acid and other aldehydes.
- Gender differences.
- Differences in deposition patterns in the respiratory tract of humans versus rodents.
- Dose-response data for DNA effects.
- Human metabolism data (qualitatively metabolism between rodents and humans appear similar, but rates may be different).

Should an RfC be developed separately? If an RfC or MOE were to be developed, which factors should be considered to account for uncertainties in the extrapolations applied?

The group agreed that developing a separate RfC is justified. The potential role of lesions such as atrophy and hyperplasia would have to be considered in the context of later tumor outcome. Uncertainty factors would include one to account for animal to human extrapolation (based on further study of the PBPK model) and one for intrahuman variability.

What mechanistic data would be useful for development of risk estimates of exposures via the oral route?

The group did not evaluate the oral exposure route but agreed that more than site-specific (i.e., URT) effects need to be examined. More data are needed to learn whether using site of toxicity dose metrics is protective of other effects.

SECTION FOUR
FINAL PLENARY SESSION

Lessons Learned and Their Applications to the Development of a Human Health Risk Assessment Framework

In efforts to integrate information deliberated throughout the two colloquia and to assist in the development of the framework, the group broadly discussed the questions listed below.

- Should a common quantitative analysis be conducted when there are commonalities among toxicities?

- In the absence of case-specific PBPK models, is there a common approach for dose adjustment for interspecies extrapolation for all responses? Does this differ for different routes of exposure?

  In the presence of PBPK models, how does MOA information influence the dose surrogate in characterizing toxicity? Can it be different for different responses?

- In the absence of BBDR models, how does MOA information influence the default approach(es) to characterize in quantitative terms the potential risk of toxicities at low levels of exposure (i.e., beyond the range of observation)? Are there common default approaches?

- The 1996 "Proposed Guidelines for Carcinogen Risk Assessment" have recommended that five factors be considered when determining the margin of exposure. These included intraspecies variation, interspecies variation, nature of the response, steepness of the dose-response curve, and biopersistence.

  The current quantitative approach for noncancer effects generally involves development of a single RfD/RfC for a "critical effect." Factors used include intraspecies variation, interspecies variation, subchronic to chronic extrapolation, LOAEL to NOAEL extrapolation, and completeness of the data base. An additional factor may be applied to account for scientific uncertainties in the study selected for derivation of the RfD/RfC.

  If the goal is to harmonize across toxicities, can a consistent set of factors be identified? How does MOA information influence the choice of these factors?

Discussions focused on criteria and factors one should consider when evaluating integrated risk assessment approaches. In addition, factors relevant to MOE application and appropriate "uncertainty" factors were detailed. Prior to these discussions, the group clarified terminology related to dose response:

Linear: When assuming a linear dose response, the ED_{10} (or point of departure) assumes that from the point of departure (POD) there is a linear extrapolation down to zero.

Nonlinear: For a nonlinear dose response, the ED_{10} (or "benchmark dose") is divided by uncertainty factors to develop an RfD.

MOE: The MOE is the ED_{10} divided by the human exposure estimate of interest. It can be applied to linear or nonlinear dose-response curves and for any endpoint.

The group agreed upon this definition of MOE but noted that the description of MOE in EPA's cancer guidelines is somewhat confusing and, therefore, needs to be clarified.

Some participants preferred the term "margin of protection;" however, it was pointed out that the term MOE was developed and used purposely so not to imply "safety" or "protection."

The group considered how adequate and useful MOE is to the risk management decision and discussed the possible basis on which an MOE should be set. The group agreed that regulators need these "numbers" for compliance purposes. Like RfDs, MOEs need to represent exposures "without appreciable risk." One participant noted that there are social, political, and legal issues as well as the science driving the decision. Another participant noted that it is ultimately a risk management decision—is the MOE acceptable given a certain set of conditions? It was noted that an MOE can be more powerful than an RfD because, in evaluating an acceptable MOE, the entire toxicity database is examined. It is the scientist/risk assessor's responsibility to bring the relevant information to the risk manager so that he/she can understand the significance of a given MOE.

Colloquium participants agreed on the following points or questions regarding the application of MOEs:

- IRIS needs to include additional risk characterization information. One participant commented that it could be included in Section 6.

- A criteria list is needed to guide risk assessors and managers in applying the MOE concept (a consistent series of questions). The list should include uncertainty issues for cancer and noncancer effects.

  One participant noted that a consistent approach may be difficult (across programs and the different regions).

- Both the numerator (ED_{10}) and denominator (human exposure of interest) values need to be clearly explained to the risk manager, including the confidence in each value.

- Adequacy of the MOE will be based largely on experience.

- Factors considered when deriving an RfD and when deciding on an MOE are similar, but not identical. While both consider toxicity and dose-response, one important distinction is that application of an MOE also considers the magnitude and uncertainty in the exposure estimate. Furthermore, as mentioned previously, the entire toxicity database is considered when deciding on an MOE.

- Mode of action needs to be carefully examined when deciding if MOE is the most scientifically viable approach for assessing risks.

The group listed the following key "uncertainty" factors for consideration when integrated approaches are applied. No "values" were assigned.

- Intraspecies differences: Differences in toxicokinetics and toxicodynamics within species.

- Interspecies differences: Differences in toxicokinetics and toxicodynamics across species.

- Quantitative linkages between toxicokinetics and toxicodynamics.

- Severity of endpoint/effects.

- Structure activity relationship information.

- Human exposure scenario information (e.g., frequency, pattern, etc.).

- Confidence limits on ED_{10} (experimental variability).

- Shape and steepness of dose-response curve.

- Integration of multiple factors.

- Species specificity/sensitivity.

- Quality of database.

- Quality of individual studies.

- Knowledge of MOA.

- Reversibility/irreversibility of effects.

- Biopersistence (e.g., is it sequestered in fat?) (toxicokinetics).

- Bioavailability (toxicokinetics).

- Particularly susceptible population (e.g., children, genetic susceptibility, pre-existing disease).
- Route of exposure.
- Route to route extrapolation.
- Relationship between MOA and human exposure scenarios.
- Confidence in PBPK models.
- Biopersistence in the environment.
- Biomarkers of effect/exposure.

Overview/Next Steps

Both colloquia were instrumental in soliciting expert opinion on evolving issues related to MOA and integrated risk assessment approaches. Participants offered their impressions on the current state of scientific knowledge and on the next steps in developing a human health risk assessment framework. Having worked through the case studies, the group agreed that, in light of available knowledge, new more scientifically-based approaches can and should be applied. The group clearly recognized, however, that many uncertainties exist. The following ideas were communicated by participants and reiterated throughout the colloquium.

- As was evidenced through case study discussions, a range of opinions still exist on the best approach (e.g., shape of the dose response curve, common MOAs, etc.).
- Before integrated risk assessment approaches can fully evolve, more quantitative information is needed.
- Risk assessors will inevitably be faced with limited data sets. The general scheme of toxicologic events may be known, but specific mechanisms may not be fully understood. What do we do if only limited MOA information is available? Do we fall back on current default approaches? Scientists will need to evaluate when "enough" data are available.
- The process requires a good deal of data interpretation. Developing a system to aid in this process will be challenging. Others agreed, asking "Can we come up with an approach that is scientifically viable and useful from a regulatory perspective?"
- As integrated approaches are explored further, a case study(ies) that would use an MOE approach needs to be developed. A set of key factors related to cancer and noncancer effects also should be formally developed.
- The overall goal of the risk assessment framework is to consider how to practice and communicate the "best science" in predicting risks.

- # The best available science should be used to generate the most credible risk assessment, but presented in a way that is useful to the risk manager.

- # Scientists need to know when not to harmonize, even when similar MOAs exist.

In closing, members of the health effects framework planning committee provided a brief overview of next steps in the framework development. The input from agency experts during this colloquia series will be reviewed. Numerous questions and issues were raised that will need to be re-examined and/or further explored. The planning committee would like to see discussions from this colloquia series expanded. A collaborative workshop, including EPA and outside groups (e.g., SOT and SRA) is being contemplated.

Participants noted that additional forums would be helpful in offering additional insight. The group also expressed interest in future colloquia to discuss topics such as exposure and health outcome data and PBPK models.